OIPE IAP87
FEB 07 2007
PATENT & TRADEMARK OFFICE

Docket No.: 1615.1019

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re the Application of:

Maxwell WELLS

Serial No. 09/556,086 Group Art Unit: 2128

Confirmation No. 7449

Filed: April 21, 2000 Examiner: Fred O. Ferris III

For: MUSIC SEARCHING METHODS BASED ON HUMAN PERCEPTION

APPEAL BRIEF

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

I. Real Party in Interest

The inventors, Maxwell WELLS, David WALLER, and Navdeep S. DHILLON, assigned all rights in the subject application to CantaMetrix, Inc. according to the Assignment executed March 28 and April 4 and 18, 2000, which was recorded at Reel 10771, Frames 620-625. CantaMetrix, Inc. subsequently assigned all rights in the application to CDDB, Inc. on May 17, 2002 according to the Assignment recorded on November 19, 2004 at Reel 16006, Frames 879-888. CDDB, Inc. was renamed Gracenote, Inc. as indicated by the documents executed June 25, 2002 which were recorded on May 13, 2004 at Reel 15341, Frames 243-262. Therefore, the real party in interest is Gracenote, Inc.

II. Related Appeals and Interferences

There are no related appeals or interferences known to Appellants, Appellants' legal representatives or the Assignee, Gracenote, Inc., which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

III. Status of Claims

Claims 1-20, 22-24, 26, 27, 29-31 and 33-43 stand rejected under 35 U.S.C. § 103(a). Claims 21, 25, 28 and 32 have been canceled.

IV. Status of Amendments

No Amendment was filed in response to the final Office Action mailed July 7, 2006.

V. Summary of Claimed Subject Matter

The application is directed to creating and using a database of musical recordings that can be searched based on human perception of the recordings as indicated by parameters extracted from the recordings. The extracted parameters are combined with a weighting assigned to each parameter to generate a single number representing a descriptor for each recording. Human listeners provide an indication of their perception of an initial set of recordings and the weightings assigned to the parameters extracted from the initial set of recordings are adjusted to match the human perception of the initial set of recordings. The adjustments to the weightings of the initial set of recordings are applied to the remaining recordings in the database. In one embodiment, subsets of the parameters are combined by formulas that correspond to terms easily recognizable by human listeners, such as energy, happiness and danceability, as described on pages 13 and 14, and the adjustments are made based on how the human listeners rank musical recordings according to each of these terms.

Claim 1

Claim 1 recites a "method for building a computational model of human perception of a descriptor of music" (claim 1, lines 1-2). Substantially all of the specification is related to techniques that can be used for such a method. Page 4, lines 13-16 of the application refers to such a computational model and modeling human perception of music is described on pages 7-9 and pages 13-15. In Fig. 3, block 305 is "model of descriptors" and the same text appears in blocks 404 and 605 in Figs. 4 and 6, respectively.

Claim 1 also recites "extracting from each of at least 5 electronic representations of musical recordings at least two numeric parameters" (claim 1, lines 3-4). Definitions of thirteen parameters that can be extracted from musical recordings are provided on pages 15-17 of the application. Parameter extractors 102, 204, 303 and 402 are illustrated in Figs. 1-4 and described on page 6, lines 30-32 ; page 7, lines 6-7; page 8, lines 7-10; and page 18, lines 26-27.

Claim 1 also recites "for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording" (claim 1, lines 5-6). Examples of weighting the extracted parameters and combining to form a single number or scalar can be found at page 8, lines 11-15 of the application and is represented in Fig. 3 by the conversion of parameters 304 to a model of descriptors 305 and similar conversions of parameters 403 and 604 to models 404 and 605 in Figs. 4 and 6. Examples of weightings are provided on page 13. A set of "descriptors (scalar)" 706 is illustrated in Fig. 7 and described at page 23, lines 19-20 as being produced by the "processes in Figure 3" (page 23, line 19).

Claim 1 also recites "adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners" (claim 1, last 3 lines). This operation is represented in Fig. 3 by model adjustment 307 which is described at page 8, lines 15-17 and an example is described at page 20, line 5 to page 9 of the application with reference to Fig. 6.

Claim 3

Claim 3 recites a "method for generating a data record associated with a music recording, the record comprising two or more scalar descriptors, each descriptor numerically describing the recording of music with which the data record is associated" (claim 3, lines 1-3). While the term "data record" is not used in the application, descriptor databases 309 and 405 are illustrated in Figs. 3 and 4 and as described at page 14, line 23 and page 18, lines 27-29 store "scalar descriptors" (page 7, line 32) for "each song" (page 14, line 22), where the descriptors relate to "the music, such as ... the song, the date of recording, etc." (page 15, lines 22-23). Also, the "databases containing recordings of music" (page 1, line 6) and "musical recording" (page 5, line 25) make it clear that the song or music can be from a recording.

Claim 3 also recites "extracting from an electronic representation of the recording of music at least two numeric parameters" (claim 3, lines 4-5). As discussed above, parameter extractors 102, 204, 303 and 402 are illustrated in Figs. 1-4 and described on page 6, lines 30-32 ; page 7, lines 6-7; page 8, lines 7-10; and page 18, lines 26-27.

Claim 3 also recites "combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording, where the weightings were previously determined" (claim 3, lines 6-8). As discussed above, examples of

weighting the extracted parameters and combining to form a single number or scalar can be found at page 8, lines 11-15 of the application and is represented in Figs. 3, 4, 6 and 7 by the parameters 304, 403 and 604 and model of descriptors 305, 404 and 605, as well as the set of "descriptors (scalar)" 706 illustrated in Fig. 7 and described at page 23, lines 19-20 as being produced by the "processes in Figure 3" (page 23, line 19).

The process of determining the weightings recited in claim 3 starts with "extracting from an electronic representation of each of at least 5 musical recordings the same at least two numeric parameters" (claim 3, lines 9-10). In the example of extracting descriptors for line dancing on page 12, five musical recordings are discussed and recitation of "at least 5 musical recordings" has been present in claim 3 since the application was filed. The overall description of the process at page 8, lines 5-20, refers to "a subset of all music" (page 8, line 19) used "to extract parameters" (page 8, line 8). Parameters are described as having a "value" e.g., page 8, line 13 and in the description of the likeness model on pages 20-21, the first step is to "subtract the parameters" (page 20, line 32), calculate the absolute differences" (page 21, line 1) and "sum the differences" (page 21, line 4). Therefore, the parameter values are numeric.

The process of determining the weightings recited in claim 3 continues with "for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording" (claim 3, lines 11-12). As noted above, a scalar or single number "descriptor 305 is created by combining the parameters with different weightings for each parameter" (page 8, lines 11-12).

The process of determining the weightings recited in claim 3 ends by "adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners" (claim 3, last 3 lines). An example of this process is described at page 8, lines 15-17.

Claim 5

Claim 5 recites a "computer readable medium containing a computer extracted data record associated with a music recording" (claim 5, lines 1-2). The use of a computer database to store data related to a recording is described at page 5, lines 17-20.

Claim 5 also recites "two or more scalar descriptors, each descriptor numerically describing the recording of music with which the data record is associated" (claim 5, lines 3-4). As discussed above with respect to claim 3, "scalar descriptors" (page 7, line 32) of a "musical

recording" (page 5, line 25) are stored in descriptor databases 309 and 405 (Figs. 3 and 4) as described at page 14, line 23 and page 18, lines 27-29.

Claim 5 also recites "extracting from an electronic representation of the recording of music at least two numeric parameters" (claim 5, lines 5-6). As discussed above, parameter extractors 102, 204, 303 and 402 are illustrated in Figs. 1-4 and described on page 6, lines 30-32 ; page 7, lines 6-7; page 8, lines 7-10; and page 18, lines 26-27.

Claim 5 also recites "combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording, where the weightings were previously determined" (claim 5, lines 7-9). As discussed above, examples of weighting the extracted parameters and combining to form a single number or scalar can be found at page 8, lines 11-15 of the application and is represented in Figs. 3, 4, 6 and 7 by the parameters 304, 403 and 604 and model of descriptors 305, 404 and 605, as well as the set of "descriptors (scalar)" 706 illustrated in Fig. 7 and described at page 23, lines 19-20 as being produced by the "processes in Figure 3" (page 23, line 19).

The process of determining the weightings recited in claim 5 starts with "extracting from an electronic representation of each of at least 5 musical recordings the same at least two numeric parameters" (claim 5, lines 10-11). As discussed above, five musical recordings are in the example of extracting descriptors for line dancing on page 12 and a "a subset of all music" (page 8, line 19) is used "to extract parameters" (page 8, line 8) on page 8. Furthermore, it is clear from, e.g., the description of the likeness model on pages 20-21, that the parameter values are numeric.

The process of determining the weightings recited in claim 5 continues with "for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording" (claim 5, lines 12-13). As noted above, a scalar or single number "descriptor 305 is created by combining the parameters with different weightings for each parameter" (page 8, lines 11-12).

The process of determining the weightings recited in claim 5 ends with "adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners" (claim 5, last 3 lines) As noted above, an example of this process is described at page 8, lines 15-17.

Claim 6

Claim 6 recites a "method for searching a database of data records associated with music recordings to find a desired recording" (claim 6, lines 1-2). From the first sentence of the application, it is clear that the invention involves searching a database of data records. As noted above, the use of a computer database to store data related to a recording is described at page 5, lines 17-20 and it is stated that the database is used to search for "desired musical recordings" (page 3, line 12).

Claim 6 also recites "identifying a comparison data record associated with a music recording in a computer readable database containing a plurality of data records, each associated with a music recording, the data records each comprising two or more scalar descriptors, each descriptor numerically describing the recording of music with which the data record is associated" (claim 6, lines 3-6). As discussed above, "scalar descriptors" (page 7, line 32) are obtained for "each song" (page 14, line 22), where the descriptors are used to describe (see, the paragraph spanning pages 3 and 4) "the music, such as ... the song, the date of recording, etc." (page 15, lines 22-23).

The process of generating each descriptor is recited in claim 6 starting with "extracting from an electronic representation of the recording of music at least two numeric parameters" (claim 6, lines 7-8). As discussed above, parameter extractors 102, 204, 303 and 402 are illustrated in Figs. 1-4 and described on page 6, lines 30-32 ; page 7, lines 6-7; page 8, lines 7-10; and page 18, lines 26-27.

The process of generating each descriptor as recited in claim 6 continues with "combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording, where the weightings were previously determined by" (claim 6, lines 10-12) As discussed above, examples of weighting the extracted parameters and combining to form a single number or scalar can be found at page 8, lines 11-15 of the application and is represented in Figs. 3, 4, 6 and 7 by the parameters 304, 403 and 604 and model of descriptors 305, 404 and 605, as well as the set of "descriptors (scalar)" 706 illustrated in Fig. 7 and described at page 23, lines 19-20 as being produced by the "processes in Figure 3" (page 23, line 19).

The process of generating each descriptor as recited in claim 6 continues with "extracting from an electronic representation of each of at least 5 musical recordings the same at least two numeric parameters" (claim 6, lines 13-14). As discussed above, five musical recordings are in

the example of extracting descriptors for line dancing on page 12 and a "a subset of all music" (page 8, line 19) is used "to extract parameters" (page 8, line 8) on page 8. Furthermore, it is clear from, e.g., the description of the likeness model on pages 20-21, that the parameter values are numeric.

The process of generating each descriptor as recited in claim 6 continues with "for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording" (claim 6, lines 15-16). As noted above, a scalar or single number "descriptor 305 is created by combining the parameters with different weightings for each parameter" (page 8, lines 11-12).

The process of generating each descriptor as recited in claim 6 ends with "adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners" (claim 6, lines 17-19). As noted above, an example of this process is described at page 8, lines 15-17.

Claim 6, ends with "searching the database to find data records with descriptors that are similar to the descriptors to the comparison record" (claim 6, lines 20-21). Searching for similar music is described from the bottom of page 19 to page 26. The use of descriptors in this process are described at page 23, line 18-22, for example, and an example of a search is provided on page 26 of the application.

Claim 10

Claim 10 recites a "method for building a computational model of human perception of likeness between musical recordings" (claim 10, lines 1-2). An example of creating a "likeness model" is described on pages 24-25.

The body of claim 10 begins by reciting "extracting from each of at least 5 electronic representations of musical recordings at least two numeric parameters" (claim 10, lines 3-4). As discussed above with respect claim 1, definitions of thirteen parameters that can be extracted from musical recordings are provided on pages 15-17 of the application. Parameter extractors 102, 204, 303 and 402 are illustrated in Figs. 1-4 and described on page 6, lines 30-32 ; page 7, lines 6-7; page 8, lines 7-10; and page 18, lines 26-27.

Claim 10 also recites "receiving from one or more human listeners who compare pairs of the musical recordings an indication of the human's perception of likeness for each compared

pair of recordings" (claim 10, lines 5-7). An example of this process is described at page 23, lines 25-30.

Claim 10 also recites "for each compared pair of the recordings, comparing each numeric parameter of one recording in the pair with the corresponding parameter of the second recording of the pair using an algorithm which produces a parameter comparison number representing the parameter comparison" (claim 10, lines 8-11). While the term "parameter comparison number" is not used in the application, it is clear that the "list of numbers ... calculated for the comparison of each song to each other song ... consist[ing] of a value for each descriptor where the value is the difference between the descriptor value for a first song and the value of the same descriptor for a second song" (page 24, lines 3-6) can be used like parameter comparison numbers.

Claim 10 next recites "for each compared pair of the recordings, combining the parameter comparison numbers with a weighting for each parameter comparison number to compute a single difference number representing the difference between the two recordings of the pair" (claim 10, lines 12-14). An example of the weighting of the differences in the list generated as described above is provided at page 24, line 20 to page 25, line 12.

Claim 10 ends by reciting "adjusting the weightings for the comparison numbers to find a set of weightings where each computed difference number for each pair of recordings most closely matches perceptions reported for the pair of recordings by the one or more human listeners" (claim 10, lines 15-17). An example applying the adjustment process described at page 8, lines 15-17 to a similarity search is provided at page 26, lines 2-10

Claim 18

Claim 18 recites a "method for creating a database of differences between music recordings" (claim 18, line 1). An example of such a database is the "multi dimensional database 607" (page 22, line 14) in Fig. 6 and a process of creating such a database is described on pages 20-26.

Claim 18 also recites "associating an identifier with each recording of a plurality of music recordings" (claim 18, line 3). Although the words "associating" and "identifier" are not used in the application, the use of identifiers in databases is known to persons of ordinary skill in the art and thus it is inherent for "songs in the target database" (page 22, line 17), where, as noted above, "songs" and "music recordings" are used substantially interchangeably in the application.

Claim 18 also recites "extracting from each recording of the plurality of recordings at least two numeric parameters" (claim 18, lines 4-5). As stated in the sentence spanning pages 19 and 20 of the application, "[t]he processes in Figure 3 are repeated to create a set of parameters 604." As discussed above, parameter extractors 102, 204, 303 and 402 are illustrated in Figs. 1-4 and described on page 6, lines 30-32 ; page 7, lines 6-7; page 8, lines 7-10; and page 18, lines 26-27 and thirteen parameters are defined on pages 15-17.

Claim 18 also recites "computing from the extracted parameters for each of a plurality of pairs of the recordings a number which represents the difference between the recordings of the pair" (claim 18, lines 6-7). An example of calculating a difference between values of a parameter extracted from different songs is provided at page 20, lines 19-25.

Claim 18 also recites "assembling the computed difference numbers into a database where each computed difference is associated with the identifier for each of the two recordings from which the difference was computed" (claim 18, lines 8-10). In an example disclosed in the application, "storing and organizing the parameter differences data is ... [provided by] a multi dimensional vector in a multi dimensional database 607" (page 22, lines 13-14) as illustrated in Fig. 6.

Claim 26

Claim 26 recites a "method for finding a music recording which is perceived by humans to be like another music recording" (claim 26, lines 1-2). As noted above, one of the objectives of the invention is "to find music that sounds to human listeners like any musical composition selected by a user" (page 19, lines 31-32).

Claim 26 also recites "receiving a specification of a target music recording" (claim 26, line 3). In the example described on page 26, "[t]he likeness database may be quickly searched by starting with any song" (page 26, line 1).

Claim 26 also recites "searching a database containing computed difference numbers between the target recording and a plurality of other recordings for those recordings which have a small computed difference number from the target music recording" (claim 26, lines 4-6). The example described on page 26 explains that what is found is "a list of likeness matches, including some that are somewhat different" (page 26, line 5), but the list generated "is much broader than the list displayed for the user" (page 26, lines 12-13), because it "includes songs that are less similar to the initial target song than would be tolerated by the listener ... [and] lie below the presentability threshold for the initial target" (page 26, lines 13-15).

VI. Grounds of Rejection to be Reviewed on Appeal

In the July 7, 2006 Office Action, the Examiner noted that claims 1-20, 22-24, 26, 27, 29-31 and 33-43 were pending in the application and rejected all of the pending claims under 35 USC § 103(a) as unpatentable over an article entitled "Music Content Analysis through Models of Audition" by Martin et al. (Reference U in the January 18, 2006 Office Action) in view of U.S. Patent 5,918,223 to Blum et al. (Reference A in the June 17, 2004 Office Action). At issue are the following:

(1) Whether Martin et al. "teaches [a] method for building a computational model of human perception of music" (Office Action, page 7, lines 17-18)?

(2) In what context does "Martin specifically set... forth that only a human listener can 'identify genre' and realize 'what other pieces or kinds of music it bears similarity to'" (Office Action, page 7, line 24 to page 8, line 1)?

(3) What in the prior art would suggest combining Martin et al. and Blum et al.?

(4) Would the combination of Martin et al. and Blum et al. teach or suggest all of the limitations recited in claims 1, 3, 5, 6 and 10?

(5) What in the cited prior art teaches or suggests "assembling the computed difference numbers into a database where each computed difference is associated with the identifier for each of the two recordings from which the difference was computed" (claim 18, last 3 lines?

(6) What in the cited prior art teaches or suggests "searching a database containing computed difference numbers between the target recording and a plurality of other recordings for those recordings which have a small computed difference number from the target music recording" (claim 26, last 3 lines)?

VII. Argument

Issue (1)

As discussed in the last full paragraph on page 11 of the Response filed April 18, 2006 by Certificate of Mailing and received by the U.S. Patent and Trademark Office (USPTO) on April 21, 2006, the rejection of the claims starts with an assertion that Martin et al. "teaches [a] method for building a computational model of human perception of music" (Office Action, page 7, lines 17-18). It is submitted that this vastly overstates what is disclosed by Martin et al. Following is the summary of the teachings of Martin et al. from the April 18, 2006 Response.

Disclosure by Martin et al.

The article by Martin et al. cited in rejecting the claims contains a discussion of the direction taken in research at the Machine Listening Group at the Massachusetts Institute of Technology Media Lab in the 1990s to recognize features of music. There are no details of how anything discussed therein was accomplished. The majority of the article is a rebuttal of other techniques that emphasize analysis of music using music theory and graduate-level music students. Instead, Martin et al. recommends attempting to model the abilities and reactions of non-experts to music in psychoacoustic experiments involving "[r]eal music, taken directly from FM radio" (page 5, lines 16-17), for example. On pages 4-6, three case studies are discussed which were reported in other articles that were not cited in the rejection, but may have disclosed what was used to begin to accomplish these objectives.

The description of the first case study in Martin et al. (speech/music discrimination) mentions "13 features that were thought to be useful discriminators" (page 4, lines 28-29) and experiments that showed "using only a subset of the features" (page 4, line 38) produced "equivalent performance from a classifier working on the full 13-dimensional feature space" (page 4, lines 37-38), that revealed some features "are sufficiently correlated that only one need be used" (page 4, lines 39-40) and "various three-dimensional classifiers performed statistically equivalently" (page 4, line 42). The only features in the first case study mentioned in Martin et al. were the following three in a "perceptual feature set" (page 4, line 43): spectral centroid variance, 4 Hz modulation energy and "pulse metric" (see page 4, line 45 to page 5, line 7).

The description of the second case study in Martin et al. (acoustic beat and tempo tracking), used "a constant-Q spectrogram, analyzing each channel for regions of sharply increasing energy, summing these regions across channel, and then calculating a phase-preserving narrowed autocorrelation to calculate tempo" (page 5, lines 18-20) or "decomposing the signal into six bands with sharply-tuned bandpass filters, and then analyzing the periodicity of each band's envelope independently" (page 5, lines 21-22). "The estimates from the multiple subbands ... [were] combined to give an overall estimate, and then the beat phase of the signal ... [was] estimated using simple heuristics" (page 5, lines 25-26).

The description of the third case study in Martin et al. (timbre classification), involving identification of musical instruments used to produce recordings, stated that "the most valuable features for source identification are related to the speed of energy buildup—as a function of frequency—during the onset of a note" (page 6, lines 12-14) which "is directly related to the Q (ratio of center frequency to bandwidth) of the nearby resonances in the system ... [and thus] is ...

related closely to ... a 'steady-state' feature of the sound" (page 6, lines 15-17). Specifically, "the log-lag correlogram" (page 6, line 25) which "encodes many salient features, including formant structure, pitch vibrato and jitter, tremolo, and onset skew ... [which] have been shown to be important for source identification by humans and for subjective judgements of timbre" (page 6, lines 28-31) was used instead of "resorting to short-time Fourier analysis, formation of sinusoidal 'tracks,' or assumptions about 'onset' and 'steady state" (page 6, lines 31-33).

Martin et al. Does Not Support Rejection

As discussed above, Martin et al. contains only a broad overview of experiments reported elsewhere. It is submitted that one of ordinary skill in the art with only the teachings in Martin et al. would be unable to build "a computational model of human perception of music" as asserted by the Examiner. All that Martin et al. provides is a direction in which to begin research on psychoacoustic analysis of music.

In the Response to Arguments section on pages 2-6 of the July 7, 2006 Office Action, the Examiner cited page 7, paragraphs 1-3 of Martin et al. as allegedly disclosing a "method for building a computational model of human perception of music" as recited in claim 1. What is stated in these paragraphs refutes the Examiner's position. According to Martin et al., "[w]e are **beginning** a project in which we will use the principles described above to construct a model of the early stages of human music perception" (Martin et al., page 7, lines 2-3, emphasis added). It is clear from this statement and Martin et al. taken as a whole that at the time Martin et al. was written the authors had not built a computational model of human perception of music and as a result, did not disclose how to do so in the cited article. All that is discussed in the first paragraph on page 7 of Martin et al. is a "first approximation of the goal of this system ... a system that can make some ... judgements about ... [a] piece of music as the human listener can" (Martin et al., page 7, lines 3-6).

There is nothing in the statements quoted at the end of the previous paragraph that suggests the only way to accomplish the goal identified in the first paragraph on page 7 of Martin et al. is "a computational model of human perception of music ... based on the perceptions reported by a human listener" (Office Action, page 7, lines 21-22). In support of this assertion, "page 5, para:4, page 7, para:1-3" of Martin et al. were cited by the Examiner. The lack of support for this assertion in first paragraph on page 7 is addressed above. Assuming that "page 5, para:4" refers to lines 11-17 on page 5 of Martin et al., this paragraph states:

> We have constructed several systems that can accurately determine tempo and locate the beat in musical signals of arbitrary polyphonic complexity and contain-

> ing arbitrary timbres (Scheirer 1997; Vercoe 1997; Scheirer 1998). The analysis is performed causally, online, and in real-time, and can be used predictively to guess when beats will occur in the future. We engaged in extensive analysis and verification of the second system, demonstrating its performance on a wide variety of musical samples and comparing it to the performance of human listeners in a short validation experiment. Real music, taken directly from FM radio, was used to validate this system and compare its performance to that of human listeners.

It is submitted that these statements do not support the Examiner's assertion that any of the systems mentioned in Martin et al. provide "a computational model of human perception of a descriptor of music" as recited in claim 1. Martin et al. states that a system was built that can "guess when beats will occur in the future" with a success rate similar to human listeners (although details of the system are not disclosed in Martin et al.). Such capability does not mean that the system in the "short validation experiment" constitutes a computational model of human perception of a descriptor of music.

Similarly, the second and third paragraphs on page 7 of Martin et al. do not support the assertion that Martin et al. disclosed a computational model of human perception of a descriptor of music. Rather than support the Examiner's assertion, the second paragraph on page 7 of Martin et al., described capabilities that the systems developed by the authors lack, i.e., "identify the genre of the music, ... pieces or kinds of music [that] it bears similarity to" (Martin et al., page 7, lines 9-10), and much more. Like the first paragraph, the third paragraph on page 7 of Martin et al. states a future objective of creating a model and does not even report on a model that had been created. After noting the skills required to perform the tasks described in paragraph 2, i.e., from "tapping along ... to ... identifying appropriate social scenarios" (Martin et al., page 7, lines 16-17), it is stated that "these skills, **if** robustly modeled, would be highly useful as the basis for constructing musical multimedia systems" (Martin et al., page 7, lines 19-20, emphasis added). This paragraph ends by stating that the authors "**wish to examine** those aspects of the music-listening process responsible for organizing the 'surface structure' of the five-second musical excerpt into a perceptual/cognitive structure that allows for other cognitive abilities to be brought to bear" (Martin et al., page 7, lines 20-22, emphasis added). Martin et al. does not state that the authors had examined the process used by human listeners to recognize music and had developed a model of that process.

For the above reasons, it is submitted that Martin et al. does not disclose a "method for building a computational model of human perception of music" as recited in claim 1. At most it indicates a desire on the part of the authors for such a model.

Issue (2)

The rejection of the claims states that "Martin specifically sets forth that only a human listener can 'identify genre' and realize 'what other pieces or kinds of music it bears similarity to'" (Office Action, page 7, line 24 to page 8, line 1) without citing where Martin et al. makes these statements. As noted in the April 18, 2006 Response, Martin et al. states that "the listener can say many interesting things about the music that are beyond our current ability to model[, including] identify[ing] the genre of the music" (Martin et al., page 7, lines 8-9). In other words, Martin et al. acknowledges a lack of capability to automatically detect genre and a desire to do so. The ability of the present invention to detect genre meets an expressed need in the state of the art.

As noted in the discussion of issue (1), the Response to Arguments in the July 7, 2006 Office Action cited paragraph 2 on page 7 of Martin et al. which contains the statements regarding capabilities of human listeners that are not met by the systems developed by the authors at the time that Martin et al. was published. As discussed above, these statements refer to a desire of the authors for a model of human perception of music that would have such capabilities, not a description of a model that had been developed by the authors. Thus, these statements do not support the Examiner's assertion that Martin et al. disclosed a "method for building a computational model of human perception of a descriptor of music" as recited in claim 1, but rather that existing systems, including those developed by the authors of Martin et al., did not provide a "model of human perception of a descriptor of music" like that recited in claim 1.

Issue (3)

As discussed in the April 18, 2006 Response, the rejection of the claims on pages 7-10 of the July 7, 2006 Office Action did not cite anything in the prior art that would suggest combining Martin et al. and Blum et al. to meet the limitations recited in the claims. On page 9 of the July 7, 2006 Office Action, the statements in Martin et al. that "only a human listener can 'identify genre' and realize 'what other pieces or kinds of music it bears similarity to'" was repeated as providing an "obvious motivation" to combine the teachings of Martin et al. and Blum et al. However, the statement in Martin et al. that such capabilities are desired would not lead a person of ordinary skill in the art to Blum et al., since Blum et al. does not provide that capability. It is only the subject application that has taught the Examiner the value of using weightings based on human perceptions of music. Nothing has been cited in the prior art that provides any suggestion of adjusting weightings or any other reason for a person of ordinary skill in the art to combine the teachings of Martin et al. and Blum et al.

On pages 5 and 6 in the Response to Arguments section of the July 7, 2006 Office Action, it was asserted that the "teachings of Martin and Blum" (Office Action, page 5, line 23 in general were sufficient to cause a person of ordinary skill in the art to combine the references without citing anything in either reference that explicitly or implicitly provides the suggestion to combine. The "test for implicit teachings" provided by *In re Kotzab*, 217 F.3d 1365, 1370, 55 USPQ2d 1313, 1317 (Fed. Cir. 2000), citing *In re Keller*, 642 F.2d 413, 425, 208 USPQ 871, 881 (CCPA 1981) and cases cited therein, is simply "what the combined teachings, knowledge of one of ordinary skill in the art, and the nature of the problem to be solved as a whole would have suggested to those of ordinary skill in the art." However, in the following sentence *In re Keller* holds that "[w]hether the Board relies on an express or an implicit showing, it must provide particular findings related thereto." The rejection of the claims on pages 7-10 of the July 7, 2006 Office Action did not provide any particular findings supporting an implicit suggestion to combine, other than the statements in the third paragraph on page 7 of Martin et al. As discussed above, it is submitted that these statements would not suggest combining the teachings of Martin et al. and Blum et al.

Next, the Response to Arguments stated that "motivation to combine is found in the recitation that modeling human perception of music 'would be highly useful as a basis for constructing musical multimedia systems'" (Office Action, page 5, lines 27-29), citing the third paragraph on page 7 of Martin et al. However, as noted above and in the April 18, 2006 Response and the November 17, 2004 Amendment (received by the USPTO on November 19, 2004), Blum et al. discloses extracting technical characteristics of audio, i.e., "amplitude (loudness), pitch, bandwidth, bass, brightness, and MEL-frequency cepstral coefficient (MFCCs)" (column 6, lines 25-27), not modeling human perception of music. Therefore, one of ordinary skill in the art would not be motivated to combine Martin et al. and Blum et al. as a result of the statements in the third paragraph on page 7 of Martin et al.

Thus, the statement that

> a skilled artisan working in this obviously competitive multimedia environment would have made an effort to become aware of what capabilities had already been developed in the market place, and hence would have been aware of, and known to seek out the relative teachings of the problem to be solved. Namely, the teachings of Martin and Blum

(Office Action, page 5, line 29 to page 6, line 2) amounts to an assertion that the Examiner can choose from the prior art any teaching that fits what is recited in the claims without any suggestion in the prior art that the teachings of the cited references should be combined. That is not the holding of *In re Kotzab* or *In re Keller* or any of the cases cited therein. As noted above,

the "test for implicit teachings" set forth by *In re Kotzab* requires "particular findings" supporting the combination and what has been cited by the Examiner does not support the combination.

Issue (4)

After acknowledging that "Martin does not explicitly disclose combining parameters to compute a descriptor or the use of parameter weighting" (Office Action, page 8, lines 4-5) it was asserted that the teachings in Blum et al., which were discussed in the Amendment filed November 17, 2004, could be combined with the alleged teachings in Martin et al. to make the claimed invention obvious. However, the overstatement of what was disclosed by Martin et al. combined with the lack of teaching in Blum et al. of a "model of human perception of a descriptor of music" (claim 1, lines 1-2) makes claim 1 patentable over Martin et al. and Blum et al., as discussed below.

As discussed in the April 18, 2006 Response and the November 17, 2004 Amendment, Blum et al. discloses a program that creates weightings according to a formula and cleans up data that does not represent a "descriptor for each recording most closely match[ing] perceptions reported for the recording by one or more human listeners" (claim 1, last 2 lines). Furthermore, Martin et al. does not disclose use of weightings for any purpose, or adjustment of any features "to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners" (e.g., claim 1, last 3 lines). Rather, as discussed previously, Martin et al. only discloses sets of features that the authors believed should be used in analyzing music. Even if a person of ordinary skill in the art were motivated to combine the teachings of Martin et al. and Blum et al., the result would be weighting of the features according to a formula, not using empirical data as recited in claim 1. Martin et al. only describes the use of human tests to validate the capabilities of the system and select features used by the system, not to adjust **how** the features are used. Thus, claim 1 patentably distinguishes over the combination of Martin et al. and Blum et al. for the reasons set forth in the April 18, 2006 Response.

The Response to Arguments in the sentence spanning pages 2 and 3 of the July 7, 2006 Office Action asserted that the second and third paragraphs on page 7 of Martin et al. disclosed "using the response of human listeners to identify musical parameters and 'classify' the music ..., i.e. report human perceptions" (emphasis in original). However, as discussed above, these paragraphs in Martin et al. only describe using the human listeners to validate a system that is not described in detail in Martin et al., not to identify or adjust musical parameters.

Next, the Response to Arguments in the July 7, 2006 Office Action asserted that the fourth paragraph on page 7 of Martin et al. "introduce[d] the concept of modeling perception by building 'statistical classifers' for evaluating musical 'properties' and making 'musical judgments'" (Office Action, page 3, lines 2-3). The fourth paragraph on page 7 of Martin et al., like the preceding paragraphs on page 7 refer to what the authors of Martin et al. hope to do in the future, rather than what had been accomplished, i.e.,

> we will build a system that segregates and groups blended musical objects—the perceptual correlates of chords—from the correlogram. We will investigate the properties of these objects in perception and build statistical classifiers which can use the objects and their properties to make musical judgments

(Martin et al., page 7, lines 23-26). Thus, like the first three paragraphs on page 7 of Martin et al., this paragraph only suggests a direction for research that might eventually lead one of ordinary skill in the art to develop what has been claimed by the Appellants. It is not sufficient, even if combined with Blum et al. to suggest what is recited in the claims.

In the second full paragraph on page 3 of the July 7, 2006 Office Action, it was asserted that column 6, lines 40-43 of Blum et al. disclosed "weighting techniques in order to emphasize perceptually important sections of musical sound" (Office Action, page 3, lines 13-14, emphasis in original). However, weighting "each trajectory's mean and standard deviation ... by the amplitude trajectory so that the perceptually important sections of the sound are emphasized" (Blum et al., column 6, lines 40-43) is not the same as "combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording" (claim 1, lines 5-6). What is described in the cited portion of column 6 in Blum et al. does not indicate any "combining" (claim 1, line 5) of the parameters, only weighting the individual "trajectories" so that perceptually significant values are given more importance than less significant values.

In the paragraph spanning pages 3 and 4 of the July 7, 2006 Office Action, it was asserted that the "statistical classifiers" mentioned in the fourth paragraph on page 7 in Martin et al. "sets forth using the response of human listeners to identify musical parameters and the concept of modeling perception" (Office Action, page 3, lines 19-20). First, as discussed above, the fourth paragraph on page 7 in Martin et al. is a description of what the authors intend to build, not a description of what they had developed. As a result, one of ordinary skill in the art at most might be motivated to try to develop what is claimed. What is described in Martin et al., including the fourth paragraph on page 7, when combined with the teachings in Blum et al., is insufficient to teach or suggest the operation recited on the last three lines of claim 1.

Furthermore, the final operation recited in claim 1 requires that "perceptions reported for **the** recording by one or more human listeners" (claim 1, last 2 lines) is used in adjusting the weightings. The statement in Martin et al. that the authors intend to "investigate the properties of these objects in perception and build statistical classifiers which can use the objects and their properties to make musical judgments" (Martin et al., page 7, lines 24-26) is much too general to suggest the specifics of what is recited on the last three lines of claim 1.

The last 3 lines of claims 3 and 5 and lines 17-19 of claim 6 all contain the limitation on the last three lines of claim 1. Therefore, it is submitted that claims 1, 3, 5 and 6, as well as claims 2, 4, 7-9 and 34-41 which depend therefrom, patentably distinguish over Martin et al. and Blum et al. for at least the reasons discussed above and in the April 18, 2006 Response.

Claim 10 recites how input from human listeners is used in "adjusting the weightings ... to find a set of weightings where each computed difference number for each pair of recordings most closely matches perceptions reported for the pair of recordings by the one or more human listeners" (claim 10, last 3 lines). At least for reasons similar to those discussed above with respect to claim 1, there is no suggestion of this operation in Martin et al. and Blum et al. Therefore, it is submitted that claim 10 and claims 11-17, 42 and 43 patentably distinguish over Martin et al. and Blum et al. for at least this reason.

Issue (5)

In the rejection of claim 18 (along with claims 7, 10-13, 19, 20 and 26) on the last two lines of page 9 and first 13 lines of page 10 in the July 7, 2006 Office Action, it is not clear where the limitation recited on the last five lines of claim 18 was allegedly disclosed by Martin et al. and Blum et al. The closest statement to these limitations that has been found in the rejection is the assertion that Blum et al. "considers the likeness (i.e. similarities) between the extracted representation of the various musical recordings" (Office Action, page 10, lines 3-4) and discloses "computing (calculate) the correlation between recorded sections (i.e. the stored numerical descriptors)" (Office Action, page 10, lines 6-7). However, nothing was cited in the rejection regarding where Blum et al. teaches or suggests specifically "computing from the extracted parameters for each of a plurality of pairs of the recordings a number which represents the difference between the recordings of the pair" (claim 18, lines 6-7) or "assembling the computed difference numbers into a database where each computed difference is associated with the identifier for each of the two recordings from which the difference was computed" (claim 18, last 3 lines).

In the Response to Arguments section of the July 7, 2006 Office Action, column 17, lines 20-55 of Blum et al. was cited in support for the assertion that the "'difference between music recordings' where 'extracted parameters' for 'pairs of recordings' [are used] would ... be rendered obvious ... since Blum ... teaches storing and retrieving sounds from a database based on similarity and the difference between sound samples" (Office Action, page 4, lines 17-21, emphasis in original). This portion of column 17 in Blum et al. states that "for sounds from the database which are similar to a certain sound file ... the vector for the sample sound file will be created. Then all database records will be measured for how similar their vector is to the sample vector" (column 17, lines 20-22, assuming line 20 starts with "It is also possible"). This is done by finding "the difference between each element of the sample vector and each element of every database vector" (column 17, lines 28-30) and "normaliz[ing] this difference" (column 17, lines 30-31). The normalization process is described in the remainder of the cited portion of column 17.

The process of finding similar sounds in the database as taught by Blum et al. continues after the cited portion of column 17 with finding a single value that represents "the distance of each database vector from the sample vector" (column 17, approximately lines 60-61), so that "the records can be sorted by distance to create an ordered list of the most similar sounds in the database to the sample sound" (column 18, lines 2-4). However, as noted above, claim 18 recites "assembling the computed difference numbers into a database where each computed difference is associated with the identifier for each of the two recordings from which the difference was computed" (claim 18, last 3 lines). Nothing was cited or found in Blum et al. or Martin et al. suggesting that the "distance" calculated as described in Blum et al. is stored in a database "with the identifier for each of the two recordings from which the difference was computed" as recited in claim 18. What has been found in Blum et al. (outside of what was cited) is using values stored in a database to calculate distances and sorting the results to find the shortest distance. Claim 18 recites a simpler process because it is not directed to finding "sounds from the database which are similar to a certain sound file" as taught by the cited portion of Blum et al., but rather calculating differences between entire recordings and storing that difference in a database using "the identifier for each of the two recordings from which the difference was computed." This is different than what is taught by Blum et al. and it is submitted that the combination of Blum et al. and Martin et al. does not teach or suggest creating a database of difference numbers as recited in claim 18. For the above reasons, it is submitted that claim 18, as well as claims 20 and 22-24 which depend therefrom, patentably distinguish over Martin et al. and Blum et al.

Issue (6)

Claim 26 recites "searching a database containing computed difference numbers between the target recording and a plurality of other recordings for those recordings which have a small computed difference number from the target music recording" (claim 26, last 3 lines). Therefore, claim 26 requires the existence of database of difference numbers like that created in claim 18. As discussed in the previous paragraph, the combination of Blum et al. and Martin et al. does not teach or suggest such a database. Sorting a set of distances calculated as described in Blum et al. to find the shortest distance is not the same as "searching a database" as recited in claim 26. For the above reasons, it is submitted that claim 26, as well as claims 27, 29-31 and 33 which depend therefrom, patentably distinguish over Martin et al. and Blum et al.

Summary of Arguments

For the reasons set forth above and in the Response filed April 18, 2006, it is submitted that claims 1-20, 22-24, 26, 27, 29-31 and 33-43 patentably distinguish over Martin et al. and Blum et al. , taken individually or in combination. Thus, it is respectfully submitted that the Examiner's final rejection of the claims is without support and, therefore, erroneous. Accordingly, the Board of Patent Appeals and Interferences is respectfully urged to so find and to reverse the Examiner's final rejection.

Enclosed is a check for the required fee of $500. Please charge any additional fee to our Deposit Account No. 19-3935.

Respectfully submitted,

STAAS & HALSEY LLP

Date: February 7, 2007

By: Richard A. Gollhofer
Richard A. Gollhofer
Registration No. 31,106

1201 New York Avenue, NW, Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501

VIII. Claims Appendix

1. A method for building a computational model of human perception of a descriptor of music, comprising:

a) extracting from each of at least 5 electronic representations of musical recordings at least two numeric parameters;

b) for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording;

c) adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners.

2. A computer readable medium containing a computer program which causes a computer to perform the method of claim 1.

3. A method for generating a data record associated with a music recording, the record comprising two or more scalar descriptors, each descriptor numerically describing the recording of music with which the data record is associated, comprising:

a) extracting from an electronic representation of the recording of music at least two numeric parameters;

b) combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording, where the weightings were previously determined by:

c) extracting from an electronic representation of each of at least 5 musical recordings the same at least two numeric parameters;

d) for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording;

e) adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners.

4. A computer readable medium containing a computer program which causes a computer to perform the method of claim 3.

5. A computer readable medium containing a computer extracted data record associated with a music recording, the record comprising:

two or more scalar descriptors, each descriptor numerically describing the recording of music with which the data record is associated, where each descriptor was generated by:

a) extracting from an electronic representation of the recording of music at least two numeric parameters;

b) combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording, where the weightings were previously determined by:

c) extracting from an electronic representation of each of at least 5 musical recordings the same at least two numeric parameters;

d) for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording;

e) adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners.

6. A method for searching a database of data records associated with music recordings to find a desired recording, comprising:

a) identifying a comparison data record associated with a music recording in a computer readable database containing a plurality of data records, each associated with a music recording, the data records each comprising two or more scalar descriptors, each descriptor numerically describing the recording of music with which the data record is associated, where each descriptor was generated by:

1) extracting from an electronic representation of the recording of music at least two numeric parameters;

2) combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording, where the weightings were previously determined by:

3) extracting from an electronic representation of each of at least 5 musical recordings the same at least two numeric parameters;

4) for each recording, combining the numeric parameters with a weighting for each parameter to compute a single number representing the descriptor for that recording;

5) adjusting the weightings for the parameters to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners; and

b) searching the database to find data records with descriptors that are similar to the descriptors to the comparison record.

7. The method of claim 6 further including, prior to searching the database, specifying that one of the descriptors of the comparison data record should be adjusted with an increase or a decrease, and the searching step is based on the descriptors of the comparison data record as adjusted.

8. A computer readable medium containing a computer program which causes a computer to perform the method of claim 6.

9. A computer readable medium containing a computer program which causes a computer to perform the method of claim 7.

10. A method for building a computational model of human perception of likeness between musical recordings, comprising:

a) extracting from each of at least 5 electronic representations of musical recordings at least two numeric parameters;

b) receiving from one or more human listeners who compare pairs of the musical recordings an indication of the human's perception of likeness for each compared pair of recordings;

c) for each compared pair of the recordings, comparing each numeric parameter of one recording in the pair with the corresponding parameter of the second recording of the pair using an algorithm which produces a parameter comparison number representing the parameter comparison;

d) for each compared pair of the recordings, combining the parameter comparison numbers with a weighting for each parameter comparison number to compute a single difference number representing the difference between the two recordings of the pair;

e) adjusting the weightings for the comparison numbers to find a set of weightings where each computed difference number for each pair of recordings most closely matches perceptions reported for the pair of recordings by the one or more human listeners.

11. The method of claim 10 where the algorithm includes subtraction of parameter values.

12. The method of claim 10 where the algorithm includes computing a correlation between parameter values.

13. The method of claim 10 where, prior to the step of comparing the numeric parameters:

a) the parameters for each recording are combined with a weighting for each parameter to compute a single number representing a descriptor for that recording, where

b) the weightings were previously determined by adjusting the weightings to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners, and

c) the descriptors are then used in the step of comparing the numeric parameters in place of the parameters.

14. A computer readable medium containing a computer program which causes a computer to perform the method of claim 10.

15. A computer readable medium containing a computer program which causes a computer to perform the method of claim 11.

16. A computer readable medium containing a computer program which causes a computer to perform the method of claim 12.

17. A computer readable medium containing a computer program which causes a computer to perform the method of claim 13.

18. A method for creating a database of differences between music recordings, comprising:

a) associating an identifier with each recording of a plurality of music recordings;

b) extracting from each recording of the plurality of recordings at least two numeric parameters;

c) computing from the extracted parameters for each of a plurality of pairs of the recordings a number which represents the difference between the recordings of the pair; and

d) assembling the computed difference numbers into a database where each computed difference is associated with the identifier for each of the two recordings from which the difference was computed.

19. The method of claim 18 where the computing step includes subtraction of parameter values.

20. The method of claim 18 where the computing step includes computing correlation between parameter values.

22. A computer readable medium containing a computer program which causes a computer to perform the method of claim 18.

23. A computer readable medium containing a computer program which causes a computer to perform the method of claim 19.

24. A computer readable medium containing a computer program which causes a computer to perform the method of claim 20.

26. A method for finding a music recording which is perceived by humans to be like another music recording, comprising:

a) receiving a specification of a target music recording; and

b) searching a database containing computed difference numbers between the target recording and a plurality of other recordings for those recordings which have a small computed difference number from the target music recording.

27. The method of claim 26 where the database is created by:

associating an identifier with each recording of a plurality of music recordings;

extracting from each recording of the plurality of recordings at least two numeric parameters selected from dynamic range, loudness, harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, tempo, sound salience and key;

computing from the extracted parameters for pairs of the recordings a number which represents the difference between the recordings of the pair; and

assembling the computed difference numbers into a database where each computed difference is associated with the identifier for each of the two recordings from which the difference was computed.

29. The method of claim 27 where the step of computing a number which represents the difference between the recordings of a pair of recordings includes the intermediate steps of:

a) combining the parameters for each recording with a weighting for each parameter to compute a single number representing a descriptor for that recording, where

b) the weightings were previously determined by adjusting the weightings to find a set of weightings where each computed descriptor for each recording most closely matches perceptions reported for the recording by one or more human listeners, and

c) the descriptors are then used in place of the parameters to compute a number which represents the difference between the recordings of the pair.

30. A computer readable medium containing a computer program which causes a computer to perform the method of claim 26.

31. A computer readable medium containing a computer program which causes a computer to perform the method of claim 27.

33. A computer readable medium containing a computer program which causes a computer to perform the method of claim 29.

34. A method as recited in claim 1, wherein the at least two numeric parameters are selected from dynamic range, loudness, harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, tempo, sound salience and key.

35. A method as recited in claim 1, wherein the at least two numeric parameters include at least one of harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, sound salience and key.

36. A method as recited in claim 3, wherein the at least two numeric parameters are selected from dynamic range, loudness, harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, tempo, sound salience and key.

37. A method as recited in claim 3, wherein the at least two numeric parameters include at least one of harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, sound salience and key.

38. A method as recited in claim 5, wherein the at least two numeric parameters are selected from dynamic range, loudness, harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, tempo, sound salience and key.

39. A method as recited in claim 5, wherein the at least two numeric parameters include at least one of harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, sound salience and key.

40. A method as recited in claim 6, wherein the at least two numeric parameters are selected from dynamic range, loudness, harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, tempo, sound salience and key.

41. A method as recited in claim 6, wherein the at least two numeric parameters include at least one of harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, sound salience and key.

42. A method as recited in claim 10, wherein the at least two numeric parameters are selected from dynamic range, loudness, harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, tempo, sound salience and key.

43. A method as recited in claim 10, wherein the at least two numeric parameters include at least one of harmonicity, rhythm strength, rhythm complexity, articulation, attack, note duration, sound salience and key.

IX. Evidence Appendix

(None)

X. Related Proceedings Appendix

(None)